INFORMATION AND CONTROL 6, 337–371 (1963)

Loss in Information Transmission through Two-way Channels*

FREDERICK JELINEK

School of Electrical Engineering, Cornell University, Ithaca, New York

LIST OF SYMBOLS

$x, \bar{x}$	Two-way channel input signals
$y, \bar{y}$	Two-way channel output signals
$p(\cdot,\cdot/\cdot,\cdot)$	Two-way channel transmission probability
$\{p(\cdot,\cdot/\cdot,\cdot)\}$	A set of symbols $p(\cdot,\cdot/\cdot,\cdot)$
Z_i^m	Sequence of length m of letters Z_i starting with Z_{i-1} and ending with Z_{i-m}
$\left.\begin{matrix} q_{X^mY^m}(x) \\ \bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x}) \end{matrix}\right\}$	Source probabilities of signal output $x(\bar{x})$ conditioned upon previous event X^m, Y^m $(\bar{X}^m, \bar{Y}^m)$
$\left.\begin{matrix} p_{\bar{x}}(\bar{y}/x) \\ \bar{p}_x(y/\bar{x}) \end{matrix}\right\}$	Restricted two-way channel transition probabilities of right (left output $\bar{y}(y)$ "caused" by the left (right) input $x(\bar{x})$ when signal $\bar{x}(x)$ was put into the right (left) terminal
$\sum_{f \ni \mathfrak{F}(X^m, Y^m) = x}$	Symbol for summation over all functions f satisfying the condition that $f(X^m, Y^m) = x$
$\ni$	Symbol for "such that."
V	Symbol for "over all" or "for all"
$E_{q,\bar{q}}[\cdot]$	Symbol for "expectation of $[\cdot]$ with respect to probability distributions $\{q_{X^mY^m}(x)\}$ and $\{\bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})\}$
$I(\bar{Y}_{n+1}^n ; X_{n+1}^n/\bar{X}_{n+1}^n)$	Symbol for "mutual information between events $\bar{Y}_{n+1}^n$ and X_{n+1}^n when event $\bar{X}_{n+1}^n$ is known to have occurred"

* This work which is a generalization of a part of a Ph.D. thesis, Department of Electrical Engineering, M.I.T., January 1962, was supported in part by the U. S. Army Signal Corps, the Air Force Office of Scientific Research, and the Office of Naval Research.

$(R, \bar{R})$	Point in cartesion plane whose first coordinate is R and second coordinate is $\bar{R}$
$\left.\begin{array}{l} P(\cdot) \\ \bar{P}(\cdot) \end{array}\right\}$	Symbol for a probability measure on the argument here represented by "$\cdot$"
*, $+$	Sometimes used as superscripts to indicate a particular choice of the underlying symbol
$\mathcal{F}, \bar{\mathcal{F}}, \mathcal{X}, \bar{\mathcal{X}}, \mathcal{Y}, \bar{\mathcal{Y}}, \mathcal{S}, \bar{\mathcal{S}}$	Script capital letters used, usually with arguments enclosed in a following parenthesis, as symbols for sets
$\mathscr{x}, \bar{\mathscr{x}}, \mathscr{y}, \bar{\mathscr{y}}, \mathscr{f}, \bar{\mathscr{f}}$	Script lower case letters used, usually with arguments enclosed in a following parenthesis, as symbols for sets
$\cup$	Symbol for union of sets
$\cap$	Symbol for intersection of sets

A discrete memoryless two-way channel is defined by a set of transmission probabilities $\{p(y, \bar{y}/x, \bar{x})\}$, where x and $\bar{x}$ are the left and right input signals, and y and $\bar{y}$ are the left and right received signals, respectively. If the transmission probabilities are restricted so that $p(y, \bar{y}/x, \bar{x}) = p_{\bar{x}}(\bar{y}/x)\bar{p}_x(y/\bar{x})$, and if we attach to the two channel terminals independent, finite memory, stationary signal sources which generate channel inputs depending on sequences of past inputs and outputs, then expressions for average information transmission rates in the left-to-right and right-to-left directions can be developed and their sum will be a simple information measure. When mutually independent messages are to be transmitted in opposite directions through the channel, it is desirable that they be encoded into sequences of strategy functions which together with the received signals constitute inputs to a transducer whose outputs are the channel input signals. The message source–encoder–transducer combinations are stochastically equivalent to signal sources whose outputs are governed by appropriate probabilities. We can interpret the transducer–channel combination as a derived two-way channel whose inputs are the strategy functions and whose outputs are the outputs of the underlying channel. Expressions for the information transmission rate through the two directions of the derived channel are developed and are compared to the expressions for the average information about outputs of the equivalent signal sources, transmitted through the underlying two-way channel. The values of the former expressions are found to be less than or equal to the values of the latter, the difference constituting a "coding information loss." A condition on the transmission probabilities enables us to define a class of lossless channels. Similarly another class is defined having the property that, regardless of the strategy code used, the information transmitted through the derived channel will be strictly less than the information transmitted through the underlying channel. The consequences of the above results on the random selection of message codes are discussed. It is shown that one can obtain the number of variables to be optimized when best random codes for lossy channels are desired, by using the number of variables for lossless channels as an exponent to the product of the size of the input and output signal alphabets. For the lossy channel class a simplified encoding procedure must in practice be applied, but as can be demonstrated, it will not yield optimal codes.

I. INTRODUCTION

A discrete memoryless two-way channel[1] shown schematically in Fig. 1 consists of two terminals, each equipped with a transmitter and a receiver. The left terminal transmits signals x from an alphabet of size

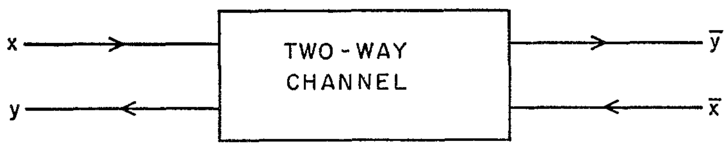

FIG. 1. Two-way channel

g and receives signals y from an alphabet of size h. The right terminal transmits signals $\bar{x}$ from an alphabet of size $\bar{g}$ and receives signals $\bar{y}$ from an alphabet of size $\bar{h}$. The channel operates synchronously: at given time intervals signals x and $\bar{x}$ are simultaneously transmitted and as a result signals y and $\bar{y}$ are received. The transmission properties of the channel are given by the probability set $\{p(y, \bar{y}/x, \bar{x})\}$. The channel is memoryless so that

$$\Pr(y_1, \cdots, y_n; \bar{y}_1, \cdots, \bar{y}_n/x_1, \cdots, x_n; \bar{x}_1, \cdots, \bar{x}_n) = \prod_{i=1}^{n} p(y_i, \bar{y}_i/x_i, \bar{x}_i) \quad (1)$$

where $\alpha_i(\alpha = y, \bar{y}, x, \bar{x})$ is a signal appearing at the terminals of the channel at time i.

If the channel is the only available medium through which the two terminals can communicate, then the most general signal source-channel arrangement is the one shown in Fig. 2. At time $i = 1, 2, \cdots$ the left

[1] The pioneering work in this area is Shannon (1961). See also Jelinek (1962a) and a summary of this work, Jelinek (1962b).

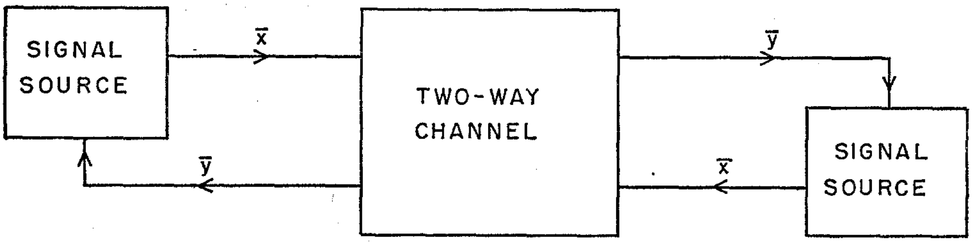

FIG. 2. Most general signal source-channel network

(right) stochastic source selects for transmission the signal $x_i(\bar{x}_i)$, depending on past received and transmitted sequences $y_1, \cdots, y_{i-1}$ and $x_1, \cdots, x_{i-1}(\bar{y}_1, \cdots, \bar{y}_{i-1}$ and $\bar{x}_1, \cdots, \bar{x}_{i-1})$. Let us impose the reasonable condition that the sources be stationary and have finite memory. Then the operation of the left source will be defined by the probability set $\{q_{X^m Y^m}(x)\}$ and that of the right source by the set $\{\bar{q}_{\bar{X}^m \bar{Y}^m}(\bar{x})\}$, where we have adopted the capital letter notation for sequences of symbols:

$$Z_i^m \equiv z_{i-1}, z_{i-2}, \cdots, z_{i-m} \tag{2}$$

(throughout (2) the symbol z stands for either of the letters x, y, $\bar{x}$, x), and where we define

$$q_{X^m Y^m}(x) \equiv \Pr(x/(x_{-1}, \cdots, x_{-m}) = X^m, (y_{-1}, \cdots, y_{-m}) = Y^m) \tag{3a}$$

$$\begin{aligned} &\bar{q}_{\bar{X}^m \bar{Y}^m}(\bar{x}) \\ &\qquad \equiv \Pr(\bar{x}/(\bar{x}_{-1}, \cdots, \bar{x}_{-m}) = \bar{X}^m, (\bar{y}_{-1}, \cdots, \bar{y}_{-m}) = \bar{Y}^m). \end{aligned} \tag{3b}$$

In (3a) and (3b) m is the assumed memory length, common to both sources.

In what follows we will limit our attention to those two-way channels whose transmission probability set is restricted by the relation[2]

$$p(y, \bar{y}/x, \bar{x}) = p_{\bar{x}}(\bar{y}/x)\bar{p}_x(y/\bar{x}), \tag{4}$$

where

$$\sum_{\bar{y}=0}^{h-1} p_{\bar{x}}(\bar{y}/x) = 1 \qquad x \in (0, 1, \cdots, g-1)$$

[2] For a thorough analysis of binary two-way channels restricted in this way see Jelinek (1964a), sections V to X.

$$\sum_{y=0}^{h-1} \bar{p}_x(y/\bar{x}) = 1 \qquad \bar{x} \in (0, 1, \cdots, \bar{g} - 1)$$

and

$$0 \leqq p_{\bar{x}}(\bar{y}/x) \leqq 1$$

$$0 \leqq \bar{p}_x(y/\bar{x}) \leqq 1$$

It will turn out that with the restriction (4) imposed we will be able in the following sections to bring into sharper focus certain important phenomena characteristic of two-way channel operation, in particular an important interpretation of a newly introduced concept of coding information loss.

The restricted two-way channel can be schematically represented by an interconnection of two oppositely oriented "alternating channels", as in Fig. 3. There the left-to-right channel can be considered an $\bar{h}$-state device, its state at time i being determined by the signal $\bar{x}_i$; the right-to-left channel can be considered an h-state device, its state at time i being determined by the signal x_i.

II. INFORMATION FLOW THROUGH THE CHANNEL

Consider the signal source–channel communication network of Fig. 4, in which the left source is characterized by an arbitrary probability set $q_{X^mY^m}(x)$ (see (3a)) and the right source by an arbitrary probability set $\bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})$ (see (3b)). In what follows we shall specify as a conven-

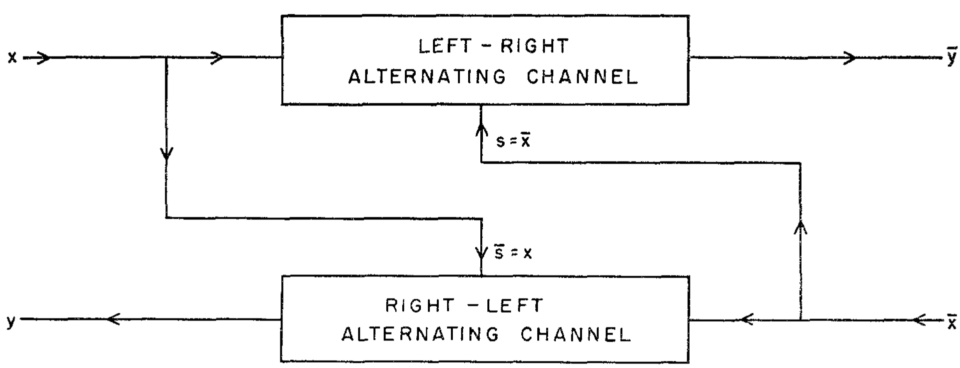

FIG. 3. Alternating channel representation of a restricted two-way channel

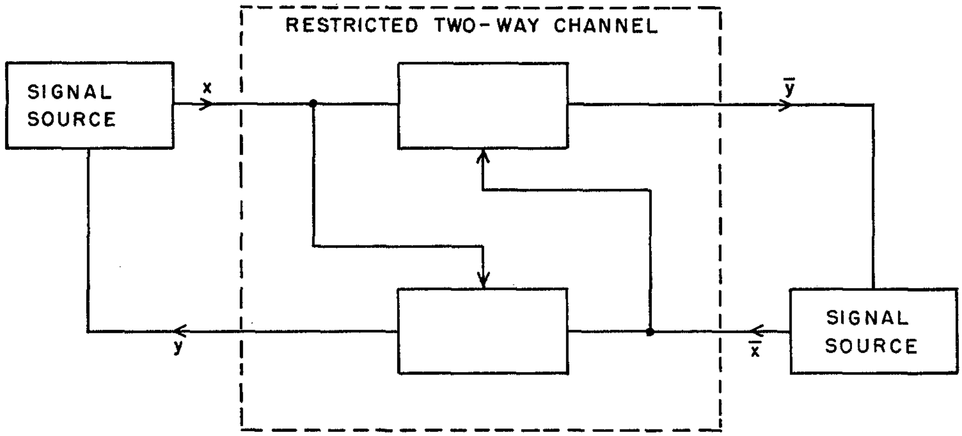

FIG. 4. Signal source-restricted channel network

tion that for purposes of signal generation the sources will behave as if

$$X_1^m = Y_1^m = \bar{X}_1^m = \bar{Y}_1^m = \underbrace{0, 0, \cdots, 0.}_{\text{sequence of length } m} \tag{5}$$

(i.e., the sequences X_1^m, Y_1^m, $\bar{X}_1^m$, and $\bar{Y}_1^m$ are composed entirely of the first letters of their respective alphabets).

We shall now derive some relations between various information measures pertaining to the two-way communication network represented in Fig. 4. We will use these later when investigating the problem of computing the capacity region.

Define the information measure

$$I'(\bar{Y}_{n+1}^n ; X_{n+1}^n/\bar{X}_{n+1}^n) = \log \frac{\Pr'(\bar{Y}_{n+1}^n/X_{n+1}^n, \bar{X}_{n+1}^n)}{\Pr'(\bar{Y}_{n+1}^n/\bar{X}_{n+1}^n)}, \tag{6}$$

where the prime indicates that the conditioning sequence $\bar{X}_{n+1}^n$ is fixed independently of the other sequences.

The expectation with respect to the given probability distributions $\{q_{X^mY^m}(x)\}$ and $\{\bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})\}$ will be denoted by $E_{q,\bar{q}}$. If all the probabilities are appropriately defined then the expectation of (6),

$$R^n(q_{X^mY^m}(x); \bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})) = E_{q,\bar{q}}[I'(\bar{Y}_{n+1}^n ; X_{n+1}^n/\bar{X}_{n+1}^n)]$$

$$= \sum_{X^n_{n+1}, \bar{X}^n_{n+1}, \bar{Y}^n_{n+1}} \Pr(X^n_{n+1}, \bar{X}^n_{n+1}, \bar{Y}^n_{n+1}) \cdot \log \frac{\Pr'(\bar{Y}^n_{n+1}/X^n_{n+1}, \bar{X}^n_{n+1})}{\Pr'\bar{Y}^n_{n+1}/\bar{X}^n_{n+1})}, \tag{7}$$

can be interpreted as the average information about the signal block transmitted from the left terminal, available at time $t = n$ at the right terminal of the channel of Fig. 4.

Note that all the information about the output of the left source is provided by the received signal sequence $\bar{Y}^n_{n+1}$, and it depends on $\bar{X}^n_{n+1}$ only in so far as the latter sequence modifies the noise in the channel. The manner of generation of $\bar{X}^n_{n+1}$ must then be irrelevant to the determination of the information transmission rate, and the latter must hence be computable from the schematic arrangement of Fig. 5. Hence the probabilities pertinent to (6) and (7) are given by (8), (10), and (11). We have

$$\Pr(X^n_{n+1}, \bar{X}^n_{n+1}, \bar{Y}^n_{n+1}) = \sum_{Y^n_{n+1}} \Pr(X^n_{n+1}, Y^n_{n+1}, \bar{X}^n_{n+1}, \bar{Y}^n_{n+1})$$

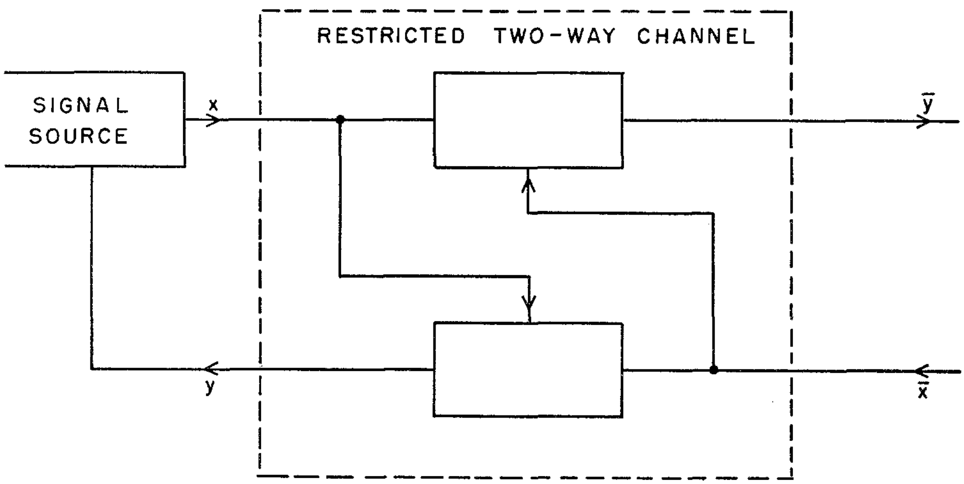

FIG. 5. Signal source-channel arrangement for computation of left-right transmission rate.

$$= \sum_{Y^n_{n+1}} \prod_{i=1}^{n} p_{\bar{x}_i}(\bar{y}_i/x_i) p_{x_i}(y_i/\bar{x}_i) q_{X_i^m Y_i^m}(x_i) \bar{q}_{\bar{X}_i^m \bar{Y}_i^m}(\bar{x}_i), \tag{8}$$

and because

$$\Pr'(X^n_{n+1}, Y^n_{n+1}, \bar{Y}^n_{n+1}/\bar{X}^n_{n+1})$$

$$= \prod_{i=1}^{n} p_{\bar{x}_i}(\bar{y}_i/x_i) \bar{p}_{x_i}(y_i/\bar{x}_i) q_{X_i^m Y_i^m}(x_i), \tag{9}$$

then

$$\Pr'(\bar{Y}^n_{n+1}/X^n_{n+1}, \bar{X}^n_{n+1})$$

$$= \frac{\sum_{Y^n_{n+1}} \Pr(X^n_{n+1}, Y^n_{n+1}, \bar{Y}^n_{n+1}/\bar{X}^n_{n+1})}{\sum_{Y^n_{n+1}, \bar{Y}^n_{n+1}} \Pr(X^n_{n+1}, Y^n_{n+1}, \bar{Y}^n_{n+1}/\bar{X}^n_{n+1})} = \prod_{i=1}^{n} p_{\bar{x}_i}(\bar{y}_i/x_i), \tag{10}$$

$$\Pr'(\bar{Y}^n_{n+1}/\bar{X}^n_{n+1}) = \sum_{X^n_{n+1}, Y^n_{n+1}} \Pr'(X^n_{n+1}, Y^n_{n+1}, \bar{Y}^n_{n+1}/\bar{X}^n_{n+1})$$

$$= \sum_{X^n_{n+1}, Y^n_{n+1}} \prod_{i=1}^{n} p_{\bar{x}_i}(\bar{y}_i/x_i) \bar{p}_{x_i}(y_i/\bar{x}_i) q_{X_i{}^m Y_i{}^m}(x_i). \tag{11}$$

In all of the preceding expressions the convention (5) is used.

In a similar way one may define the expression for the average information $\bar{R}^n(q_{X^m Y^m}(x); \bar{q}_{\bar{X}^m \bar{Y}^m}(\bar{x}))$ transmitted through the channel in the right-to-left direction by a signal block of length n. It is:

$$\bar{R}^n(q_{X^m Y^m}(x); \bar{q}_{\bar{X}^m \bar{Y}^m}(\bar{x})) = E_{q,\bar{q}}[I''(Y^n_{n+1}; \bar{X}^n_{n+1}/X^n_{n+1})], \tag{12}$$

where the double prime indicates that the conditioning sequence X^n_{n+1} is fixed independently of the other sequences.

The probabilities necessary for the determination of expression (12) can be obtained from those for expression (7) by "barring" all of the unbarred and by removing the bars from all the barred quantities in expressions (8), (9), (10), and (11), and by replacing the prime by a double prime.

For added clarity we rewrite (7) and (12) in a form which places in evidence their relationship with the channel and the sources of Fig. 4:

$$R^n(q_{X^m Y^m}(x), \bar{q}_{\bar{X}^m \bar{Y}^m}(\bar{x}))$$

$$= E_{q,\bar{q}}\left[\log \frac{\prod_{i=1}^{n} p_{\bar{x}_i}(\bar{y}_i/x_i)}{\sum_{X^n_{n+1}, Y^n_{n+1}} \prod_{i=1}^{n} p_{\bar{x}_i}(\bar{y}_i/x_i) \bar{p}_{x_i}(y_i/\bar{x}_i) q_{\bar{X}_i{}^m \bar{Y}_i{}^m}(x)}\right], \tag{13a}$$

$$\bar{R}^n(q_{X^mY^m}(x), \bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x}))$$
$$= E_{q,\bar{q}}\left[\log \frac{\prod_{i=1}^n \bar{p}_{x_i}(y_i/\bar{x}_i)}{\sum_{X^n_{n+1},\bar{Y}^n_{n+1}} \prod_{i=1}^n \bar{p}_{x_i}(y_i/\bar{x}_i)p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{q}_{\bar{X}_i{}^m\bar{Y}_i{}^m}(\bar{x})}\right]. \tag{13b}$$

The expectations are to be carried out in (13a) with the help of (8), and in (13b) with the help of a probability corresponding to (8) in the manner described in the preceding paragraph.

We shall next prove a theorem and its more important corrollary dealing with the relationships between certain information measures and above rates R^n and $\bar{R}^n$. We define

$$I(X^n_{n+1}, Y^n_{n+1}; \bar{X}^n_{n+1}, \bar{Y}^n_{n+1}) = \log \frac{\Pr(X^n_{n+1}, Y^n_{n+1}, \bar{X}^n_{n+1}, \bar{Y}^n_{n+1})}{\Pr(X^n_{n+1}, Y^n_{n+1}) \Pr(\bar{X}^n_{n+1}, \bar{Y}^n_{n+1})}. \tag{14}$$

and state

THEOREM 1. *The relationship*

$$I'(\bar{Y}^n_{n+1}; X^n_{n+1}/\bar{X}^n_{n+1}) + I''(Y^n_{n+1}; \bar{X}^n_{n+1}/X^n_{n+1}) = I(X^n_{n+1}, Y^n_{n+1}; \bar{X}^n_{n+1}, \bar{Y}^n_{n+1}) \tag{15}$$

holds, where the primes are to be interpreted as in (6) *and* (12). *Therefore the sum of simultaneous information flows in the left-to-right and right-to-left directions caused by transmission of signal blocks of length n through the restricted channel of Fig. 4 is given by the information measure* (14).

PROOF: From (8) we get:

$$\frac{\Pr(X^n_{n+1}, Y^n_{n+1}, \bar{X}^n_{n+1}, \bar{Y}^n_{n+1})}{\Pr(X^n_{n+1}, Y^n_{n+1}) \Pr(\bar{X}^n_{n+1}, \bar{Y}^n_{n+1})}$$
$$= \frac{\prod_{i=1}^n p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i)q_{X_i{}^mY_i{}^m}(x_i)\bar{q}_{\bar{X}_i{}^m\bar{Y}_i{}^m}(\bar{x}_i)}{\left[\begin{array}{l}\left[\prod_{i=1}^n q_{X_i{}^mY_i{}^m}(x_i) \sum_{\bar{X}^n_{n+1},\bar{Y}^n_{n+1}} \prod_{i=1}^n p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i)\bar{q}_{\bar{X}_i{}^m\bar{Y}_i{}^m}(\bar{x}_i)\right] \\ \cdot\left[\prod_{i=1}^n \bar{q}_{\bar{X}_i{}^m\bar{Y}_i{}^m}(x_i) \sum_{X^n_{n+1},Y^n_{n+1}} \prod_{i=1}^n p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i)q_{X_i{}^mY_i{}^m}(x_i)\right]\end{array}\right]} \tag{16}$$

Hence

$$\frac{\Pr(X_{n+1}^n, Y_{n+1}^n, \bar{X}_{n+1}^n, \bar{Y}_{n+1}^n)}{\Pr(X_{n+1}^n, Y_{n+1}^n)\Pr(\bar{X}_{n+1}^n, Y_{n+1}^n)}$$
$$= \left[\frac{\prod_{i=1}^n p_{\bar{x}_i}(\bar{y}_i/x_i)}{\sum_{\bar{X}_{n+1}^n, \bar{Y}_{n+1}^n} \prod_{i=1}^n \bar{p}_{\bar{x}_i}(\bar{y}_i/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i) q_{X_i{}^m Y_i{}^m}(x_i)}\right] \tag{17}$$
$$\cdot \left[\frac{\prod_{i=1}^n p_{\bar{x}_i}(\bar{y}_i/x_i)}{\sum_{X_{n+1}^n, Y_{n+1}^n} \prod_{i=1}^n \bar{p}_{x_i}(y_i/\bar{x}_i) p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{q}_{\bar{X}_i{}^m \bar{Y}_i{}^m}(\bar{x}_i)}\right].$$

Comparing the right hand side of (17) with those of (13a) and (13b) and bearing in mind the definitions (6) and (12), the equality (15) follows if we take logarithms on both sides of (17). Q.E.D.

COROLLARY. *For the restricted two-way channel of Fig. 4 the following relationship between information measures holds*

$$I[X_{n+1}^n, Y_{n+1}^n; \bar{X}_{n+1}^n, \bar{Y}_{n+1}^n] = I[X_{n+1}^n; \bar{X}_{n+1}^n] + I[\bar{Y}_{n+1}^n; X_{n+1}^n/\bar{X}_{n+1}^n] + I[Y_{n+1}^n; \bar{X}_{n+1}^n/X_{n+1}^n], \tag{18}$$

and therefore R^n and $\bar{R}^n$ satisfy the following relationship:

$$R^n(q_{X^mY^m}(x), \bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})) + \bar{R}^n(q_{X^mY^m}(x), \bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})) = E_{q,\bar{q}}[I(\bar{Y}_{n+1}^n; X_{n+1}^n/\bar{X}_{n+1}^n)] + E_{q,\bar{q}}[I(Y_{n+1}^n; \bar{X}_{n+1}^n/X_{n+1}^n)] + E_{q,\bar{q}}[I(X_{n+1}^n; \bar{X}_{n+1}^n)]. \tag{19}$$

PROOF: If (18) is proven, relation (19) will follow directly from (7), (12), and (15). By elementary properties of the information measure[3] we get

$$I(X_{n+1}^n, Y_{n+1}^n; \bar{X}_{n+1}^n, \bar{Y}_{n+1}^n) = I(Y_{n+1}^n; \bar{Y}_{n+1}^n/X_{n+1}^n, \bar{X}_{n+1}^n) + I(\bar{Y}_{n+1}^n; X_{n+1}^n/\bar{X}_{n+1}^n) + I(Y_{n+1}^n; \bar{X}_{n+1}^n/X_{n+1}^n) + I(X_{n+1}^n; \bar{X}_{n+1}^n). \tag{20}$$

Relation (18) will follow if we can demonstrate that the first term on the right hand side of (20) is identically zero. But

$$I(Y_{n+1}^n; \bar{Y}_{n+1}^n/X_{n+1}^n, \bar{X}_{n+1}^n) = \log \frac{\Pr(Y_{n+1}^n, \bar{Y}_{n+1}^n/X_{n+1}^n, \bar{X}_{n+1}^n)}{\Pr(Y_{n+1}^n/X_{n+1}^n, \bar{X}_{n+1}^n)\Pr(\bar{Y}_{n+1}^n/X_{n+1}^n, \bar{X}_{n+1}^n)}, \tag{21}$$

[3] See Fano (1961).

where from (8)

$$\Pr(Y_{n+1}^n, \bar{Y}_{n+1}^n / X_{n+1}^n, \bar{X}_{n+1}^n)$$
$$= \frac{\prod_{i=1}^n p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i)q_{X_i{}^mY_i{}^m}(x_i)\bar{q}_{\bar{X}_i{}^m\bar{Y}_i{}^m}(\bar{x}_i)}{\sum_{Y_{n+1}^n\bar{Y}_{n+1}^n}\prod_{i=1}^n p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i)q_{X_i{}^mY_i{}^m}(x_i)\bar{q}_{\bar{X}_i{}^m\bar{Y}_i{}^m}(\bar{x}_i)}. \tag{22}$$

Hence

$$\frac{\Pr(Y_{n+1}^n, \bar{Y}_{n+1}^n/X_{n+1}^n, \bar{X}_{n+1}^n)}{\Pr(Y_{n+1}^n/X_{n+1}^n, \bar{X}_{n+1}^n)\Pr(\bar{Y}_{n+1}^n/X_{n+1}^n, \bar{X}_{n+1}^n)}$$
$$= \frac{\sum_{Y_{n+1}^n,\bar{Y}_{n+1}^n}\prod_{i=1}^n p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i)q_{X_i{}^mY_i{}^m}(x_i)\bar{q}_{\bar{X}_i{}^m\bar{Y}_i{}^m}(\bar{x}_i)}{\left[\sum_{\bar{Y}_{n+1}^n}\prod_{i=1}^n p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{q}_{\bar{X}_i{}^m\bar{Y}_i{}^m}(\bar{x}_i)\sum_{Y_{n+1}^n}\prod_{i=1}^n \bar{p}_{x_i}(y_i/\bar{x}_i)q_{X^mY^m}(x_i)\right]} = 1 \tag{23}$$

The logarithm of the left hand side of (23) is thus equal to zero as asserted. Q.E.D.

It ought to be noted that relation (19) confirms the correctness of the interpretation of the expressions (7) and (12) as information rates: the average mutual information function $E_{q,\bar{q}}[I(X_{n+1}^n; \bar{X}_{n+1}^n)]$ has nonzero value whenever communication between the channel terminals is established causing the signal sources to be correlated. Thus the quantities $E_{q,\bar{q}}[I(\bar{Y}_{n-1}^n; X_{n-1}^n/\bar{X}_{n-1}^n)]$ and $E_{q,\bar{q}}[I(Y_{n-1}^n; \bar{X}_{n-1}^n/X_{n-1}^n)]$ do not account for the total information transmission through the channel.

Theorem 1 will prove to be helpful in what follows. We shall show in Section IV that if for a given channel we denote by K^m the set of all points $(R, \bar{R})$ of information transmission rates obtainable for some block length n by some source probability assignments $\{q_{X^mY^m}(x)\}$ and $\{\bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})\}$, then unlike in the one-way channel case, it will not be possible in general to find codes signalling through the given two-way channel with error probability tending to zero at rates approaching arbitrarily closely the boundaries of the region K^m. It turns out that not all of the information (7) and (12) is usable for decoding of messages.

It ought to be stressed that without restriction (4) it is impossible to separate the oppositely directed information flows from one another, as was done above. The simplicity and interpretability of the results of Section IV will also depend directly on restriction (4).

III. CODING AND SOURCES WITH FINITE MEMORY

Let the memory length m of the sources in Fig. 4 be given and suppose that it is possible to find, for a nonnegative number λ, and for some integer m, those (possibly not unique) probability sets $\{q_{X^mY^m}(x_i)\}$ and $\{\bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})\}$ which would maximize the sum

$$R^n(q_{X^mY^m}(x), \bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})) + \lambda\bar{R}^n(q_{X^mY^m}(x), \bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})). \tag{24}$$

Consider now the problem of transmitting through the channel at the rates R^n and $\bar{R}^n$, obtained from the maximization of the expression (24), information generated by two completely independent stationary stochastic sources, located one at each terminal. Past results of information theory would immediately suggest that the independent sources be encoded so that the resulting transmitted signal sequences would have the statistical properties of the optimized sources in Fig. 4. That is, considering without any loss of generality the left terminal, one would attempt to encode the messages into strings of symbols that would then serve as inputs to some transducer. This transducer, possibly with the help of signals y received in the past, would put out signals x. One would then wish to adjust the entire message generation-encoding-transducing process so as to make it statistically describable by the optimal set $\{q_{X^mY^m}(x)\}$. A diagram of this scheme is provided in Fig. 6.

We will now describe briefly a suitable transducer and its input alphabet.[4] Let us define functions f of "memory" length m mapping the space of sequence pair X^m, Y^m on the space of channel input signals x:

$$f(X^m, Y^m) = x. \tag{25}$$

Similar functions are defined for the right terminal.

A function f is fully defined if a table of values for its $(gh)^m$ different possible arguments is given. Thus it can be represented by a g-nary sequence of $(gh)^m$ elements, each corresponding to a different point in the domain of definition. We can write

$$f \equiv (c_0, c_1, \cdots, c_i, \cdots, c_{(gh)^m-1}) \tag{26}$$

where c_i is the value of f when the sequence X^mY^m constituting its

[4] For a more thorough discussion of the transducer see Jelinek (1964a), section IV.

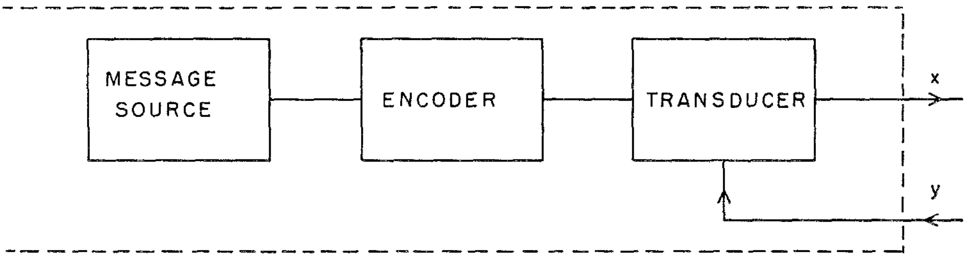

FIG. 6. Equivalent finite memory signal source derived from an independent memoryless message source.

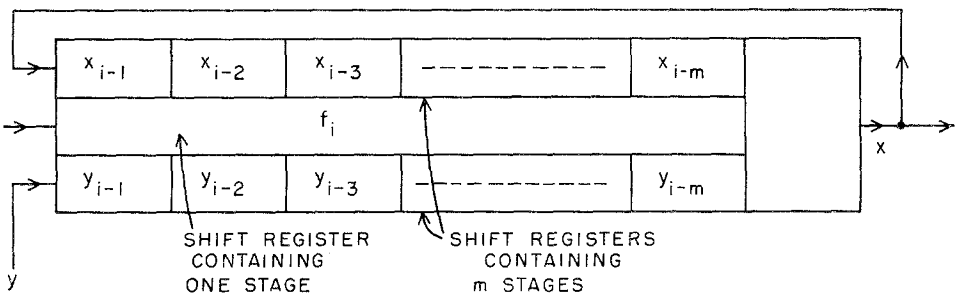

FIG. 7. Finite state function-signal transducer

argument is the g-nary representation of the integer i. It is then clear from (26) that there are altogether $g^{(gh)^m}$ different possible functions f.

Consider next the transducer of Fig. 7 whose outputs could constitute the signal inputs to a two-way channel. The transducer is a device consisting of a top and bottom shift register of m stages, and a middle shift register of one stage. At time i the state of the transducer determines the output x_i and is itself determined by the contents $x_{i-1}, \cdots, x_{i-m}$ of the top register, $y_{i-1}, \cdots, y_{i-m}$ of the bottom register, and f_i of the middle register. In the next time interval all of the register contents are shifted to the right by one stage, the contents of the rightmost stages (that is, x_{i-m}, f_i, y_{i-m}) being eliminated, and the leftmost stages are filled from top to bottom by the symbols x_i, f_{i+1}, y_i. The state at time $i + 1$ and the output signal x_{i+1} are now determined, and a new cycle may begin. The number of stages m corresponds to the memory length of the symbols f.

We may now take a pair of such transducers, attach them to the two-way channel and connect their inputs to stationary stochastic

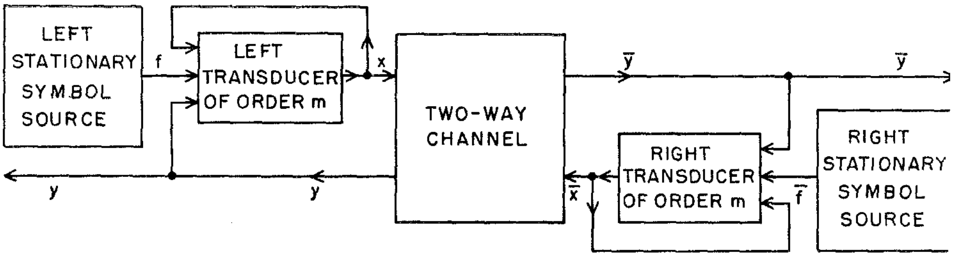

FIG. 8. Symbol source–transducer–two-way channel network

sources generating symbols f and $\bar{f}$ synchronously at given time intervals. Such a source–transducer–two-way channel communication network is schematically represented in Fig. 8. The source-transducer combinations have the effect of the signal sources in Fig. 4. In fact, it can be shown[5] that if the sources are independent of each other and generate successive symbols f and $\bar{f}$ independently with arbitrary probabilities $P(f)$ and $\bar{P}(f)$, respectively, then the equivalent signal sources resulting from the source-transducer combinations generate signals with probabilities $q_{X^mY^m}(x)$ and $\bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})$, respectively. The latter probabilities can be computed from the former by use of the expressions[5]

$$\begin{aligned} q_{X^mY^m}(x) &= \sum_{f \ni f(X^m,Y^m)=x} P(f) \\ \bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x}) &= \sum_{\bar{f} \ni \bar{f}(\bar{X}^m,\bar{Y}^m)=\bar{x}} \bar{P}(\bar{f}) \end{aligned} \tag{27}$$

It may be of interest to note that both the summations involve in general $g^{(gh)^{m-1}}$ terms.

It is known[6] that sequences of outputs of any ergodic message source can be encoded into sequences of symbols of the f-alphabet so that the output behavior of the message source–encoder combination would approximate that of a source generating successive symbols f independently with arbitrary prescribed probabilities $P(f)$.

The encoding problem stated in the paragraph following Eq. (24) is thus reduced to the problem of finding sets of probabilities $\{P(f)\}$ and $\{\bar{P}(\bar{f})\}$ governing the symbol generation of the sources in Fig. 8 which would make the source–transducer combinations generate signals with prescribed probabilities $\{q_{X^mY^m}(x)\}$ and $\{\bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})\}$.

[5] See Theorem 1 of Jelinek (1964*a*).
[6] See Fano (1961), chap. 6, pp. 181–214.

The author demonstrated in another paper[7] that appropriate probability sets $\{P(f)\}$ and $\{\bar{P}(\bar{f})\}$ can be found, and he developed a straightforward construction procedure for accomplishing this. It may be helpful to restate here the pertinent theorem.

THEOREM 2. *Let g be the size of the input signal alphabet x, and h the size of the output alphabet y. Then $(g-1)(gh)^m + 1$ is the maximum number of nonzero probabilities $P(f)$ needed to satisfy the equations*

$$q_{X^m Y^m}(x) = \sum_{\forall f \ni f(X^m, Y^m) = x} P(f) \tag{28}$$

for all X^m, Y^m, x for any arbitrary set $q_{X^m Y^m}(x)$ whose elements have the property that

$$\sum_{x=0}^{g-1} q_{X^m Y^m}(x) = 1 \qquad q_{X^m Y^m}(x) \geqq 0 \tag{29}$$

for all X^m, Y^m, x.

It should be pointed out that the number of functions f (or $\bar{f}$) which are assigned a nonzero probability ought to be as small as possible in order to make the message encoding least complicated. We are here faced with a large problem indeed since there are $g^{(gh)^m}$ different possible functions f (if $g = h = 2$ and $m = 2$ the number of functions f is 65,536). The number of probabilities $q_{X^m Y^m}(x)$ in a set is "only" $g(gh)^m$ (note: $2(2.2)^2 = 32$) and it is remarkable that the number of needed non-zero probabilities $P(f)$ turns out to be at most $(q-1)$ $(gh)^m + 1$.

IV. INFORMATION LOSS DUE TO CODING

Suppose that function sources are attached to transducers connected to a given two-way channel, as in Fig. 8. Let the sources be independent from one another and let them generate successive symbols f and $\bar{f}$ with probabilities $P^*(f)$ and $\bar{P}^*(\bar{f})$, respectively. Define signal source probabilities as in (27), and define

$$S^n(P(f); \bar{P}(\bar{f})) \equiv E_{P(f),\bar{P}(\bar{f})}[I(Y^n_{n+1}; F^n_{n+1}/\bar{F}^n_{n+1})]$$

$$= \sum_{F^n_{n+1}, \bar{F}^n_{n+1}, \bar{Y}^n_{n+1}} \Pr(F^n_{n+1}, \bar{F}^n_{n+1}, \bar{Y}^n_{n+1}) \log \frac{\Pr(\bar{Y}^n_{n+1}/F^n_{n+1}, \bar{F}^n_{n+1})}{\Pr(\bar{Y}^n_{n+1}/\bar{F}^n_{n+1})} \tag{30}$$

[7] This work is found in Jelinek (1964*b*). A special case of the problem restricted to binary two-way channels is treated in Jelinek (1964*a*), section IV, Theorem 2.

and

$$\bar{S}^n(P(f); \bar{P}(\bar{f})) \equiv E_{P(f)\bar{P}(\bar{f})}\,[I(Y^n_{n+1}\,; \bar{F}^n_{n+1}/F^n_{n+1})]$$

$$= \sum_{F^n_{n+1},\, Y^n_{n+1},\, \bar{F}^n_{n+1}} \Pr(F^n_{n+1}, Y^n_{n+1}, \bar{F}^n_{n+1}) \log \frac{\Pr(Y^n_{n+1}/F^n_{n+1}, \bar{F}^n_{n+1})}{\Pr(Y^n_{n+1}/F^n_{n+1})}. \quad (31)$$

In (27), (30), (31) and all equations to follow, notation (2) is being used, and convention (5) is adhered to. Since F^n_{n+1} together with Y^n_{n+1} completely specify X^n_{n+1}, and $\bar{F}^n_{n+1}$ together with $\bar{Y}^n_{n+1}$ completely specify $\bar{X}^n_{n+1}$, the quantity S^n is the average information about the left source in Fig. 8 available at the right channel terminals after a block of n letters f has been transmitted; similarly, $\bar{S}^n$ is the average information about the right source in Fig. 8 available at the left channel terminals after a block of n letters $\bar{f}$ has been transmitted. The question we intend to ask is: What is the relationship between $S^n(P^*(f), \bar{P}^*(\bar{f}))$ and $R^n(q^*_{X^nY^n}(x), \bar{q}^*_{\bar{X}^n\bar{Y}^n}(\bar{x}))$ (see definition (7)), and between $\bar{S}^n(P^*(f), \bar{P}^*(\bar{f}))$ and $\bar{R}^n(q^*_{X^nY^n}(x), \bar{q}^*_{\bar{X}^n\bar{Y}^n}(\bar{x}))$ (see definition (12))?

Before we state and prove our results, we should like to remark that by a random coding argument it can be demonstrated that given any sets $\{P(f)\}$ and $\{\bar{P}(\bar{f})\}$ and any $\epsilon > 0$ and a positive integer n, there exist codes signalling at rates within ϵ of the point

$$[S^n(P(f), (\bar{P}(\bar{f})), \bar{S}^n(P(f), \bar{P}(\bar{f}))] \quad (32)$$

and for these codes the probability of decoding error decreases to zero exponentially with increasing code-word length.[8] In view of the fact that there are $g(gh)^m$ members of the set $\{q_{X^mY^m}(x)\}$ and $g^{(gh)^m}$ members of the set $\{P(f)\}$, the answer to the question asked in the preceding paragraph is important for message coding. For should it turn out that $S^n = R^n$ and $\bar{S}^n = \bar{R}^n$, then whenever trying to maximize the sum $S^n + \lambda\bar{S}^n$ over all possible assignments $\{P(f)\}$ and $\{\bar{P}(\bar{f})\}$, one would maximize the sum $R^n + \lambda\bar{R}^n$ over all possible assignments $\{q_{X^mY^m}(x)\}$ and $\{\bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})\}$ instead, and then would find the appropriate sets $\{P(f)\}$ and $\{\bar{P}(\bar{f})\}$ by the construction procedure of Jelinek (1964 b). Unfortunately, as will be seen, $S^n = R^n$ and $\bar{S}^n = \bar{R}^n$ only in the restricted case of so called noiseless channels (see section 4.3), otherwise always $S^n < \bar{S}^n < \bar{R}^n$.

[8] See Jelinek (1962*a*), Theorem 6.2 and its proof in section 6, pp. 46–80. These and additional results will be submitted for publication in the near future.

A. RELATIONS BETWEEN THE RATES $(S^n, \bar{S}^n)$ AND $(R^n, \bar{R}^n)$

THEOREM 3. *Let a pair of independent sources generating successive symbols f and $\bar{f}$ independently with probabilities $P^*(f)$ and $\bar{P}^*(f)$ be attached to the transducers of a two-way channel, as in Fig. 8. If the probabilities $\{q^*_{X^m Y^m}(x)\}$ and $\{\bar{q}^*_{\bar{X}^m \bar{Y}^m}(\bar{x})\}$ are defined by (27), then*

$$R^n(q^*_{X^m Y^m}(x), \bar{q}^*_{\bar{X}^m\bar{Y}^m}(\bar{x})) + \bar{R}^n(q^*_{X^m Y^m}(x), \bar{q}^*_{\bar{X}^m\bar{Y}^m}(\bar{x})) = E_{P^*(f),\bar{P}^*(\bar{f})}[I(F^n_{n+1}, Y^n_{n+1}; \bar{F}^n_{n+1}, \bar{Y}^n_{n+1})]. \quad (33)$$

PROOF: We have

$$E_{P^*(f)\bar{P}^*(\bar{f})}[I(F^n_{n+1}, Y^n_{n+1}; \bar{F}^n_{n+1}, \bar{Y}^n_{n+1})] = \sum_{\substack{F^n_{n+1}, Y^n_{n+1} \\ \bar{F}^n_{n+1}, \bar{Y}^n_{n+1}}} \Pr(F^n_{n+1}, Y^n_{n+1}, \bar{F}^n_{n+1}, \bar{Y}^n_{n+1}) \log \cdot \frac{\Pr(F^n_{n+1}, Y^n_{n+1}, \bar{F}^n_{n+1}, \bar{Y}^n_{n+1})}{\Pr(F^n_{n+1}, Y^n_{n+1}) \Pr(\bar{F}^n_{n+1}, \bar{Y}^n_{n+1})} \quad (34)$$

In view of Theorem 1 (see Eq. (15)) and definitions (7) and (12), relation (33) will follow if we can show that

$$E_{P^*(f),\bar{P}^*(\bar{f})}[I(F^n_{n+1}, Y^n_{n+1}; \bar{F}^n_{n+1}, \bar{Y}^n_{n+1})] = E_{q^*,\bar{q}^*}[I(X^n_{n+1}, Y^n_{n+1}; \bar{F}^n_{n+1}, \bar{Y}^n_{n+1})]. \quad (35)$$

It should be noted that because of the convention (5), the sequence $F^n_{n+1}(\bar{F}^n_{n+1})$ is a function mapping the sequence $Y^n_{n+1}(\bar{Y}^n_{n+1})$ into a sequence $X^n_{n+1}(X^n_{n+1})$. The relations

$$P(F^n_{n+1}) = \prod_{i=1}^{n} P(f_i)$$
$$\bar{P}(\bar{F}^n_{n+1}) = \prod_{i=1}^{n} P(\bar{f}_i) \quad (36)$$

hold and therefore

$$\sum_{F^n_{n+1} \ni F^n_{n+1}(Y^n_{n+1}) = X^n_{n+1}} P(F^n_{n+1}) = \Big(\sum_{f \ni f(X_n{}^m, Y_n{}^m) = x_n} P(f)\Big)\Big(\sum_{f \ni f(X^m_{n-1}, Y^m_{n-1}) = x_{n-1}} P(f)\Big) \cdots \cdots \Big(\sum_{f \ni f(X_1{}^m, Y_1{}^m) = x_1} P(f)\Big).$$

Thus by (27),

$$\sum_{F_{n+1}^n \ni F_{n+1}^n (Y_{n+1}^n) = X_{n+1}^n} P(F_{n+1}^n) = \prod_{i=1}^n q_{X_i{}^m Y_i{}^m}(x_i)$$
$$\sum_{\bar{F}_{n+1}^n \ni \bar{F}_{n+1}^n (\bar{Y}_{n+1}^n) = \bar{X}_{n+1}^n} \bar{P}(F_{n+1}^n) = \prod_{i=1}^n \bar{q}_{\bar{X}_i{}^m \bar{Y}_i{}^m}(\bar{x}). \tag{38}$$

We now proceed to prove relation (35). It will be helpful to simplify the notation in what follows. Instead of Z_{n+1}^n we will simply write Z (see (2)).

It is easily seen that

$$\Pr\ (Y, \bar{Y}, F, \bar{F}) = \prod_{i=1}^n p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i)P(f_i)\bar{P}(\bar{f}_i),$$

where

$$\begin{aligned} x_i &= f_i(X_i{}^m, Y_i{}^m) \\ \bar{x}_i &= \bar{f}_i(\bar{X}_i{}^m, \bar{Y}_i{}^m) \end{aligned} \qquad \text{for } i = 1, 2, \cdots, n. \tag{39}$$

Hence

$$\Pr\ (Y, \bar{Y}, F, \bar{F}) = \Pr'\ (Y/F(Y), \bar{F}(\bar{Y}))\ \Pr''\ (\bar{Y}/F(Y), \bar{F}(\bar{Y}))P(F)\bar{P}(\bar{F}), \tag{40}$$

where $\Pr'\ (\bar{Y}/X, \bar{X})$ is as in (10), and $\Pr''\ (\bar{Y}/X, \bar{X})$ has a corresponding form. Also

$$\begin{aligned} \Pr(Y, F) &= \sum_{\bar{Y}, \bar{F}} \Pr(Y, \bar{Y}, F, \bar{F}) \\ &= \sum_{\bar{X}} \Pr''(Y/F(Y), \bar{X}) \sum_{\bar{Y}} \Pr'(\bar{Y}/F(Y), \bar{X}) \\ &\qquad \sum_{\bar{F} \ni \bar{F}(\bar{Y}) = \bar{X}} P(F)\bar{P}(\bar{F}). \end{aligned} \tag{41}$$

It follows from (10), (40), and (41) that for given sequences Y and $\bar{Y}$ the measure $\log[\Pr(Y, \bar{Y}, F, \bar{F})/\Pr(Y, F)\Pr(\bar{Y}, \bar{F})]$ will have the same value for all functions F and $\bar{F}$ such that $F(Y) = X$ and $\bar{F}(\bar{Y}) = \bar{X}$, where X and $\bar{X}$ are some fixed sequences. Thus, after some rather obvious cancellations and substitutions we can write, using (38) and (10),

$$
\begin{aligned}
\sum_{Y,\bar{Y},F,\bar{F}} \Pr(Y, \bar{Y}, F, \bar{F}) \log \frac{\Pr\,(Y, \bar{Y}, F, \bar{F})}{\Pr\,(Y, F)\,\Pr\,(\bar{Y}, \bar{F})} \\
= \sum_{X,\bar{X},Y,\bar{Y}} \prod_{i=1}^{n} p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i) q_{X_i{}^m Y_i{}^m}(x_i)\bar{q}_{\bar{X}_i{}^m \bar{Y}_i{}^m}(\bar{x}_i) \\
\cdot \left(\log \frac{\prod_{i=1}^{n} p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i)}{\left\{ \begin{array}{l} [\sum_{\bar{X}',\bar{Y}'} \prod_{i=1}^{n} p_{\bar{x}_i}(\bar{y}_i'/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i')\bar{q}_{\bar{X}_i^{m\prime}\bar{Y}_i^{m\prime}}(\bar{x}_i{}')] \\ \cdot [\sum_{X',Y'} \prod_{i=1}^{n} p_{\bar{x}_i}(\bar{y}_i/x_i')\bar{p}_{x_i}(y_i'/\bar{x}_i) q_{X_i^{m\prime}Y_i^{m\prime}}(x_i{}')] \end{array} \right\}} \right),
\end{aligned}
\tag{42}
$$

Comparing the left hand side of (42) with (34), and the right hand side with (17), equality (35) follows. Q.E.D.

Theorem 3 shows that in Fig. 8 the average mutual information between blocks of symbols (Y_{n+1}^n, F_{n+1}^n) and $(\bar{Y}_{n+1}^n, \bar{F}_{n+1}^n)$ at opposing terminals of the derived channel (see Fig. 9) is equal to the sum of average signal information rates transmitted in the opposite directions through the two-way channel (see the discussion at the beginning of Section II). Since, as postulated, the sources in Fig. 8 generate functions f and $\bar{f}$ independently, S^n, the average information about the left source transmitted through the derived channel, is equal to the average mutual information between $(\bar{Y}_{n+1}^n, \bar{F}_{n+1}^n)$ and F_{n+1}^n, and similarly, $\bar{S}^n$, the average information about the right source transmitted through the derived channel, is equal to the average mutual information between (Y_{n+1}^n, F_{n+1}^n) and $\bar{F}_{n+1}^n$. Thus Theorem 3 indicates that

$$R^n + \bar{R}^n \geqq S^n + \bar{S}^n,$$

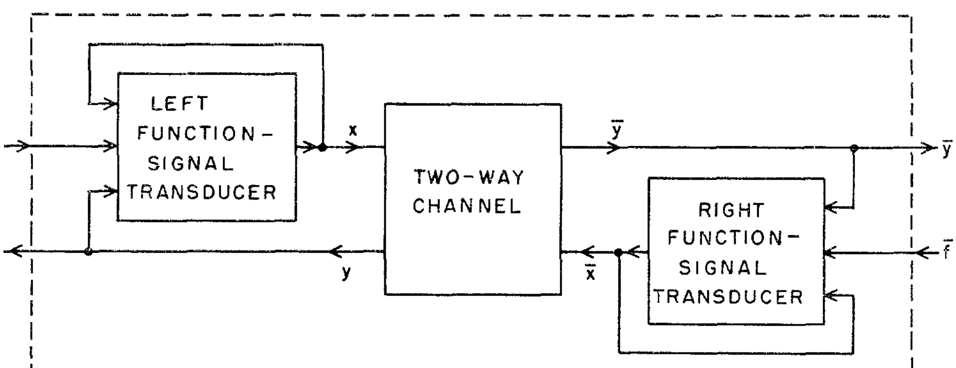

Fig. 9. Derived two-way channel

and indeed from elementary information measure relations[3] we have

$$E[I(F, Y; \bar{F}, \bar{Y})] = E[I(\bar{Y}; F/\bar{F})] + E[I(Y; \bar{F}/F)] \\ + E[I(Y; \bar{Y}/F, \bar{F})] + E[I(F; \bar{F})], \tag{43}$$

and our conjecture follows from (30), (31), and (33), since

$$E[I(F; \bar{F})] = 0,$$

and

$$E[I(Y; \bar{Y}/F, \bar{F})] \geqq 0. \tag{44}$$

We shall first sharpen the above result and then study its significance.

THEOREM 4. *Under the assumptions of Theorem* 3,

$$R^n(q^*_{X^mY^m}(x), \bar{q}^*_{\bar{X}^m\bar{Y}^m}(\bar{x})) - S^n(P^*(f), \bar{P}^*(\bar{f})) \\ = E_{P^*,\bar{P}^*}[I'(Y^n_{n+1}; \bar{Y}^n_{n+1}/F^n_{n+1}, \bar{X}^n_{n+1})] \geqq 0, \tag{45}$$

equality on the right hand side holding if and only if, whenever $P(F^n_{n+1}) \neq 0$ *and* $\bar{P}(\bar{F}^n_{n+1}) \neq 0$ *then*

$$\mathrm{Pr}'(\bar{Y}^n_{n+1}/F^n_{n+1}(Y^n_{n+1}), \bar{F}^n_{n+1}(\bar{Y}^n_{n+1})) = K(\bar{Y}^n_{n+1}, F^n_{n+1}, \bar{F}^n_{n+1}) \tag{46}$$

over all Y^n_{n+1} *such that*

$$\mathrm{Pr}''(Y^n_{n+1}/F^n_{n+1}(Y^n_{n+1}), \bar{F}^n_{n+1}(\bar{Y}^n_{n+1})) \neq 0.$$

The prime signs in (45) and (46) are to be interpreted as in (6) and (12).

PROOF: In what follows we will again use the simplified notation (43). We will first prove the middle equality part of (45). It will follow if we can show that

$$I'(\bar{Y}; X/\bar{X}) - I(\bar{Y}; F/\bar{F}) - I'(Y; \bar{Y}/\bar{F}, \bar{X}) = 0 \tag{47}$$

whenever

$$F(Y) = X \text{ and } \bar{F}(\bar{Y}) = \bar{X}.$$

We have

$$I'(\bar{Y}; X/\bar{X}) = \log \frac{\mathrm{Pr}'(\bar{Y}/X, \bar{X})}{\mathrm{Pr}'(\bar{Y}/\bar{X})},$$

$$I(\bar{Y}; F/\bar{F}) = \log \frac{\mathrm{Pr}(\bar{Y}/F, \bar{F})}{\mathrm{Pr}(\bar{Y}/\bar{F})}, \tag{48}$$

$$I'(Y; \bar{Y}/F, \bar{X}) = \log \frac{\mathrm{Pr}'(\bar{Y}/F(Y), \bar{X})}{\mathrm{Pr}'(\bar{Y}/F, \bar{X})},$$

where the probabilities $\Pr'(\bar{Y}/X, \bar{X})$ and $\Pr'(\bar{Y}/\bar{X})$ can be obtained from (10) and (11), respectively. Also from (41) it follows that if $\bar{F}(\bar{Y}) = \bar{X}$, then

$$\Pr(\bar{Y}/\bar{F}) = \sum_X \Pr'(\bar{Y}/X, \bar{F}(\bar{Y})) \sum_Y \Pr''(Y/X, \bar{F}(\bar{Y})) \sum_{F \ni F(Y)=X} P(F) \tag{49}$$

$$= \sum_{X,Y} \prod_{i=1}^{n} p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i) q_{X^m Y^m}(x_i)$$

where the last equality was obtained with help of (10) and (38). Thus from (11), whenever $\bar{F}(\bar{Y}) = \bar{X}$, then

$$\Pr(\bar{Y}/\bar{F}) = \Pr'(\bar{Y}/\bar{X}). \tag{50}$$

Also from (40),

$$\Pr(\bar{Y}/F, \bar{F}) = \frac{\sum_Y \Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y}))\ \Pr''(Y/F(Y), \bar{F}(\bar{Y}))P(F)\bar{P}(\bar{F})}{\sum_{Y,\bar{Y}} \Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y}))\ \Pr''(Y/F(Y), \bar{F}(\bar{Y}))P(F)\bar{P}(\bar{F})}. \tag{51}$$

But from (39),

$$\sum_{Y,\bar{Y}} \Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y}))\Pr''(Y/F(Y), \bar{F}(\bar{Y}))$$

$$= \sum_{\substack{y_1,\cdots,y_{n-1}\\ \bar{y}_1,\cdots,\bar{y}_{n-1}}} \prod_{i=1}^{n-1} p_{\bar{x}_i}(\bar{y}_i/x_i)\bar{p}_{x_i}(y_i/\bar{x}_i) \cdot \tag{52}$$

$$\cdot \sum_{y_n,\bar{y}_n} p_{\bar{x}_n}(\bar{y}_n/x_n)\bar{p}_{x_n}(y_n/\bar{x}_n),$$

where

$$\begin{aligned} x_i &= f_i(X_i^m, Y_i^m), \\ \bar{x}_i &= \bar{f}_i(\bar{X}_i^m, \bar{Y}_i^m) \end{aligned} \qquad i = 1, \cdots, n,$$

and the sum over y_n, $\bar{y}_n$ on the right hand side of (52) is equal to unity since neither x_n nor $\bar{x}_n$ depend on y_n, $\bar{y}_n$. Thus by recursion of this argument we get

$$\sum_{Y,\bar{Y}} \Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y}))\Pr''(Y/F(Y), \bar{F}(\bar{Y})) = 1 \tag{53}$$

for all F, $\bar{F}$.

Hence, after an obvious cancellation, we get from (51)

$$\begin{aligned}\Pr(\bar{Y}/F, \bar{F}) &= \sum_{Y} \Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y}))\Pr''(Y/F(Y), \bar{F}(\bar{Y})) \\ &= \sum_{Y} \Pr'(\bar{Y}/F(Y), \bar{X})\Pr''(Y/F(Y), \bar{X}),\end{aligned} \tag{54}$$

if $\bar{F}(\bar{Y}) = \bar{X}$. Moreover,

$$\Pr'(\bar{Y}/F, \bar{X}) = \frac{\sum_{Y} \Pr'(\bar{Y}/F(Y), \bar{X}) \Pr''(Y/F(Y), \bar{X})P(F)}{\sum_{Y,\bar{Y}} \Pr(\bar{Y}/F(Y), \bar{X}) \Pr''(Y/F(Y), \bar{X})P(F)}. \tag{55}$$

It can be shown by an argument similar to the one leading to (53) that the denominator on the right hand side of (55) is equal to 1. Thus by comparison with (6),

$$\Pr(\bar{Y}/F, \bar{F}) = \Pr'(\bar{Y}/F, \bar{X}) \tag{56}$$

if $\bar{F}(\bar{Y}) = \bar{X}$. Since also

$$\Pr'(\bar{Y}/F(Y), \bar{X}) = \Pr'(\bar{Y}/X, \bar{X}) \tag{57}$$

if $F(Y) = X$, then from (48), (50), (56), and (57) the equality (47) follows.

We must now show that

$$E_{P^*,\bar{P}^*}[I'(Y; \bar{Y}/F, \bar{X})] \geqq 0, \tag{58}$$

and that equality is possible if and only if condition (46) is fulfilled. The negative of the left hand side of (58) is from (54) and (56) equal to

$$\sum_{Y,F,\bar{Y},\bar{F}} \Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y})) \Pr''(Y/F(Y), \bar{F}(\bar{Y}))P(F)\bar{P}(\bar{F}) \cdot \left(\log \frac{\sum_{Y} \Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y})) \Pr''(Y/F(Y), \bar{F}(\bar{Y}))}{\Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y}))}\right) \tag{59}$$

$$\leqq \sum_{\bar{Y},\bar{F},Y,F} \Pr''(Y/F(Y), \bar{F}(\bar{Y}))P(F)\bar{P}(\bar{F}) \Pr'(\bar{Y}/F, \bar{F}) - 1 = 0,$$

equality holding if and only if

$$\log \frac{\sum_{Y} \Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y})) \Pr''(Y/F(Y), \bar{F}(\bar{Y}))}{\Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y}))} = 0, \tag{60}$$

whenever

$$\Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y}))\Pr''(Y/F(Y), \bar{F}(\bar{Y}))P(F)\bar{P}(\bar{F}) \neq 0.$$

Assuming that $P(F) \neq 0$ and $\bar{P}(\bar{F}) \neq 0$, then the required condition is:

$$\sum_{Y'} \Pr'(\bar{Y}/F(Y'), \bar{F}(\bar{Y})\Pr''(Y'/F(Y'), \bar{F}(\bar{Y})) \\ = \Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y})) \tag{61}$$

whenever

$$\Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y}))\Pr''(Y/F(Y), \bar{F}(\bar{Y})) \neq 0 \tag{62}$$

But the left hand side of (61) is independent of Y, thus under the condition (62) the right hand side must be also. Hence (46) is a necessary condition for equality in (59). The condition is also sufficient, since

$$\sum_{Y'} \Pr' \;(\bar{Y}/F(Y'), \bar{F}(\bar{Y}))\; \Pr'' \;(Y'/F(Y'), \bar{F}(\bar{Y})) \\ = K(\bar{Y}, F, \bar{F}) \sum_{Y'} \Pr'' \;(Y'/F(Y'), \bar{F}(\bar{Y}) = K(\bar{Y}, F, \bar{F}). \tag{63}$$

This completes the proof of the theorem. Q.E.D.

The relationship between the rates $\bar{R}^n$ and $\bar{S}^n$ must abstractly be the same as that between R^n and S^n. Hence we have the

COROLLARY. *Under the assumptions of Theorem* 3,

$$\bar{R}^n(q^*_{X^m Y^m}(x), \bar{q}^*_{X^m Y^m}(\bar{x})) - \bar{S}^n(P^*(f), \bar{P}^*(\bar{f})) \\ = E_{P^*,\bar{P}^*}\{I''(\bar{Y}^n_{n+1}\,;\, Y^n_{n+1}/\bar{F}^n_{n+1}, X^n_{n+1})\} \geqq 0. \tag{64}$$

equality on the right hand side holding if and only if whenever $P(F) \neq 0$ *and* $\bar{P}(\bar{F}) \neq 0$ *then*

$$\Pr'' \;(Y^n_{n+1}/F^n_{n+1}(Y^n_{n+1}), \bar{F}^n_{n+1}(\bar{Y}^n_{n+1})) = \bar{K}(Y^n_{n+1}, F^n_{n+1}, \bar{F}^n_{n+1}) \tag{65}$$

over all $\bar{Y}^n_{n+1}$ *such that*

$$\Pr' \;(\bar{Y}^n_{n+1}/F^n_{n+1}(Y^n_{n+1}), \bar{F}^n_{n+1}(\bar{Y}^n_{n+1})) \neq 0.$$

The double prime sign in (65) is to be interpreted as in (12).

B. CODING INFORMATION LOSS

We have seen in Theorem 4 that $E\{I'(Y;\, \bar{Y}/F, \bar{X})\}$ is the difference between the average information transmitted through the channel in the left-to-right direction and that part of it which is useful at the right terminal for identification of independent messages generated at the left terminal. We can thus interpret the quantity $E\{I'(Y;\, \bar{Y}/F,$

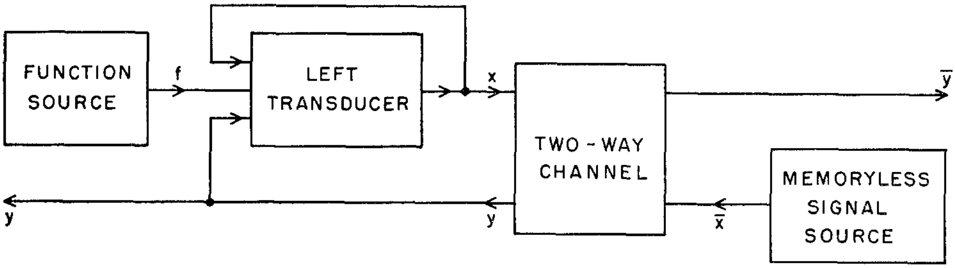

FIG. 10. Function source–channel–signal source arrangement for computation of information loss in the left transducer.

$\bar{X})\}$ as the average information lost in the left function-signal transducer (see Figs. 6 and 7). $E\{I'(Y;\bar{Y}/F,\bar{X})\}$ can be interpreted as the average information, given a transmitted sequence $\bar{X}$, that the sequence $\bar{Y}$ can provide about the sequence Y when the sequence F has already been decoded (for illustration see Fig. 10). The averaging is done by use of the probability

$$\begin{aligned}\Pr\,&(\bar{X},\bar{Y},F,Y)\\&=\Pr''\,(Y/X,\bar{F}(\bar{Y}))\,\Pr'\,(\bar{Y}/X,\bar{F}(\bar{Y}))\sum_{F\ni F(Y)=X}P(F)\end{aligned}\tag{66}$$

which gives the "actual" relative frequency of the event $\bar{X}$, $\bar{Y}$, F, Y in the ensemble. The quantity $E\{I'(Y;\bar{Y}/F,\bar{X})\}$ can rightfully be termed a loss, at least for the purposes of the right decoder which is not interested in the identification of the sequence Y after it identified F.

For similar reasons $E\{I''(\bar{Y};Y/F,X)\}$ can be interpreted as the average information loss in the right function-signal transducer.

The signal sequences that are simultaneously transmitted in opposite direction represent mutually independent messages. However, we wish the signals to be correlated in order to exploit the statistical properties of the channel so as to be able to transmit information with arbitrarily small decoding error at rates exceeding Shannon's inner bound to the capacity region, G_I. In the general channel, such attempts result in a coding loss which would be eliminated if the messages were encoded directly into signals X and $\bar{X}$ (see (46)).

C. Conditions for Absence of Coding Loss

We will now inquire what restrictions are imposed by the necessary and sufficient conditions (46) and (65) for the absence of coding loss.

We will first show that there exist special, so called noiseless channels for which (46) and (65) are satisfied regardless of the source probability sets $\{P(f)\}$ and $\{\bar{P}(\bar{f})\}$. Afterward we shall ask whether, given an arbitrary channel, sets $\{P(f)\}$ and $\{\bar{P}(\bar{f})\}$ can be found which would insure coding losslessness and yet would not be equivalent to sets selecting transmitted signals without regard to previously received signals.

THEOREM 5. *The necessary and sufficient condition* (46) *for the absence of coding loss in the left transducer is satisfied for any probability distribution over input function sequences* F_{n+1}^{n}, $\bar{F}_{n+1}^{n}$ *for those channels and those channels only whose transmission probabilities satisfy either of the following properties:*

(a) *For any signal combination* x, $\bar{x}$, *the probability* $\bar{p}_x(y/\bar{x})$ *is equal to either zero or one.*

(b) *The probability* $p_{\bar{x}}(\bar{y}/x)$ *is a function of the signals* $\bar{x}$ *and* $\bar{y}$ *only.*

PROOF: It ought to be noted that the present theorem is not restricted to sources which generate successive symbols f or $\bar{f}$ independently. We will first show that (46) is satisfied whenever either of properties (a) or (b) is. If condition (a) is satisfied then given F, $\bar{F}$ and $\bar{Y}$ there is one and only one sequence Y which can be received, as long as convention (5) is adhered to. Thus (46) holds regardless of the distribution $\{P(F)\}$ and $\{\bar{P}(\bar{F})\}$. If condition (b) holds then $\Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y}))$ is independent of $F(Y)$ and hence à fortiriori of Y, so that (46) holds again. It ought to be noted that if condition (b) holds then no communication in the left-to-right direction is possible and we are dealing with a one-way channel. If condition (a) is satisfied then the right-to-left direction is termed "noiseless" since the signals transmitted fully determine the signal received at the left terminal.

We will next show that if neither of the conditions (a) or (b) is satisfied, there will exist a combination of function sequences F and $\bar{F}$ for which (46) will fail to hold for at least some $\bar{Y}$. Thus if for such F, $\bar{F}$ the probabilities $P(F) \neq 0$ and $\bar{P}(\bar{F}) \neq 0$, the coding loss in the left transducer will be positive.

Since condition (a) does not hold, there will exist signals x^* and $\bar{x}^*$ such that $0 < \bar{p}_{x^*}(y^1/\bar{x}^*) < 1$ and $0 < \bar{p}_{x^*}(y^2/\bar{x}^*) < 1$ for some y^1 and y^2. Since condition (b) does not hold either, then there exist signals $\bar{x}^+$ and $\bar{y}^+$ such that $p_{\bar{x}^+}(\bar{y}^+/x^1) \neq p_{\bar{x}^+}(\bar{y}^+/x^2)$ for some x^1 and x^2. Consider the functions

$$\begin{aligned} \bar{F}^*(\bar{Y}) &= (\bar{x}_1 = \bar{x}^*, \bar{x}_2 = \bar{x}^+, \bar{x}_3, \cdots, \bar{x}_n) \text{ for all } \bar{Y} \\ F^*(Y) &= (x_1 = x^*, x_2 = f^*(y_1), x_3, \cdots, x_n) \text{ for all } Y, \end{aligned} \tag{67}$$

where

$$f^*(y^1) = x^1$$
$$f^*(y^2) = x^2,$$

and $(x_3, \cdots, x_n)$ and $(\bar{x}_3, \cdots, \bar{x}_n)$ are arbitrary but fixed sequences.

Let

$$\begin{aligned} Y^1 &= (y_1 = y^1, y_2, y_3, \cdots, y_n) \\ Y^2 &= (y_1 = y^2, y_2', y_3, \cdots, y_n) \end{aligned} \tag{68}$$

where $(y_3, \cdots, y_n)$ is any fixed sequence such that

$$\bar{p}_{x_i}(y_i/x_i) \neq 0 \text{ for all } i = 3, 4, \cdots, n$$

and y_2 and y_2' are such that

$$\bar{p}_{x^1}(y_2/\bar{x}^+) \neq 0$$
$$\bar{p}_{x^2}(y_2'/\bar{x}^+) \neq 0.$$

Finally, let

$$\bar{Y}^* = (\bar{y}_1, \bar{y}_2 = \bar{y}^+, \bar{y}_3, \cdots, \bar{y}_n) \tag{69}$$

where $(\bar{y}_1, \bar{y}_3, \cdots, \bar{y}_n)$ is an arbitrary but fixed sequence such that

$$p_{\bar{x}^*}(y_1/x^*) \neq 0$$
$$p_{\bar{x}_i}(\bar{y}_i/x_i) \neq 0 \qquad i = 3, 4, \cdots, n.$$

It now follows that

$$\Pr'(\bar{Y}^*/F^*(Y^1), \bar{F}^*(\bar{Y}^*)) \neq \Pr'(\bar{Y}^*/F^*(Y^2), \bar{F}(\bar{Y}^*)) \tag{70}$$

although

$$\begin{aligned} \Pr''(Y^1/F^*(Y^1), \bar{F}^*(\bar{Y}^*)) &\neq 0 \\ \Pr''(Y^2/F^*(Y^2), \bar{F}^*(\bar{Y}^*)) &\neq 0. \end{aligned} \tag{71}$$

This proves our assertion. Q E.D.

Having proven Theorem 5, we would like to turn our problem around, so to speak, and inquire whether for channels satisfying neither condition (a) nor (b) of Theorem 5 codes could be found which would not map messages strictly into sequences of input signals, and which would yet be lossless. We will show that in general this will be impossible.

THEOREM 6. *For any code of any word length n which does not map left terminal messages strictly into sequences of input signals, the coding loss*

$$E[I(Y_{n+1}^n, \bar{Y}_{n+1}^n/F_{n+1}^n, \bar{X}_{n+1}^n)] > 0 \tag{72}$$

for all two-way channels defined so that:

(a) *There is for no $\bar{x} \in (0, \cdots, \bar{g} - 1)$ any pair of signals $x^1, x^2 (x^1 \neq x^2)$ such that*

$$p_{\bar{x}}(\bar{y}/x^1) = p_x(\bar{y}/x^2) \text{ for all } \bar{y} \in (1, \cdots, n).$$

(b) *There is for no $x \in (0, \cdots, g - 1)$ any pair of signals $\bar{x}^1, \bar{x}^2$ such that whenever $\bar{p}_x(y/\bar{x}^1) \neq 0$ then $\bar{p}_x(y/\bar{x}^2) = 0$ and whenever $\bar{p}_x(y/\bar{x}^2) \neq 0$ then $\bar{p}_x(y/\bar{x}^1) = 0$.*

(c) *For any signal combination $x, \bar{x}$ there exist at least two signals $y^1 \neq y^2$ such that $\bar{p}_x(y^1/\bar{x}) \neq 0, \bar{p}_x(y^2/\bar{x}) \neq 0$.*

PROOF: We will show that if conditions (a) and (b) above are met, then for all sets of code words $\{F\} = \mathfrak{F}$ and $\{\bar{F}\} = \bar{\mathfrak{F}}$, inequality (72) will hold, unless given any $F \in \bar{\mathfrak{F}}$, $F(Y) = X$ for all Y that are receivable when F and any $\bar{F} \in \mathfrak{F}$ were transmitted. It will be shown further that if set $\mathfrak{F}$ has the latter character, then the code maps messages strictly into sequences X.

Given any F and $\bar{F}(\bar{Y}) = \bar{X}$, let $\mathcal{Y}(F, \bar{X})$ be the set of all sequences Y such that

$$\Pr''\ (Y/F(Y), \bar{X}) \neq 0. \tag{73}$$

Now the set $\mathcal{Y}(F, \bar{X})$ must contain at least 2^n different sequences Y, otherwise condition (c) could not be satisfied. If $Y^1, Y^2 \in \mathcal{Y}(F, \bar{X})$, then losslessness is possible only if

$$\Pr'\ (\bar{Y}/F(Y^1), \bar{X}) = \Pr'\ (\bar{Y}/F(Y^2), \bar{X}) \tag{74}$$

Now if $F(Y^j) = (x_1^j, x_2^j, \cdots, x_n^j)$, $j = 1, 2$ then (74) implies that

$$\prod_{i=1}^n p_{\bar{x}_i}(\bar{y}_i/x_i^1) = \prod_{i=1}^n p_{\bar{x}_i}(\bar{y}_i/x_i^2). \tag{75}$$

Assume first that the above is not equal to zero. Then there exists a constant k, $k \neq 0$ such that k, $p_{\bar{x}_n}(\bar{y}_n/x_n^1) = p_{\bar{x}_n}(\bar{y}_n/x_n^2)$, and since given F and $\bar{F}$ the choice of $\bar{x}_n$ and x_n^j is not influenced by $\bar{y}_n$, we must have $k = 1$; in fact, (46) can be satisfied only if $kp_{\bar{x}_n}(\bar{y}/x_n^{\,1}) = p_{\bar{x}_n}(\bar{y}/x_n^{\,2})$ for *all* $\bar{y} \in (1, \cdots, n)$. Thus we must always have

$$\prod_{i=1}^{n-1} p_{\bar{x}_i}(\bar{y}_i/x_i^{\,1}) = \prod_{i=1}^{n-1} p_{\bar{x}_i}(\bar{y}_i/x_i^{\,2}) \neq 0 \tag{76}$$

which, by the above argument, forces

$$p_{\bar{x}_{n-1}}(\bar{y}_{n-1}/x_{n-1}) = p_{\bar{x}_{n-1}}(\bar{y}_{n-1}/x_{n-1}^{2}).$$

Continuing in this way we see that provided the expressions in (74) are not equal to zero, the condition (46) can be satisfied only if

$$p_{\bar{x}_i}(\bar{y}_i/x_i^{\,1}) = p_{\bar{x}_i}(\bar{y}_i/x_i^{\,2}) \qquad \text{for all } i = 1, \cdots, n.$$

On the other hand, if expressions (74) are equal to zero, let m be the lowest integer such that $p_{\bar{x}_m}(\bar{y}_m/x_m^{\,1}) = 0$ and let l be the lowest integer such that $p_{\bar{x}_l}(\bar{y}_l/x_l^{\,2}) = 0$. Let $m \neq l$, and assume without loss of generality that $m < l$. But then $\prod_{i=1}^{m} p_{\bar{x}_i}(\bar{y}_i/x_i^{\,2}) \neq 0$ and there surely exists a sequence $\bar{Y}^* = (\bar{y}_1, \cdots, \bar{y}_m, \bar{y}_{m+1}^*, \cdots, \bar{y}_n^{\,*})$ and sequences $Y^{1*} = (y_1^{\,1}, \cdots, y_m^{\,1}, y_{m+1}^{1*}, \cdots, y_n^{1*})$ and $Y^{2*} = (y_1^{\,2}, \cdots, y_m^{\,2}, y_{m+1}^{2*}, \cdots, y_n^{2*})$ such that $Y^{1*}, Y^{2*} \in y(F, \bar{F}(\bar{Y}^*))$, and $Pr'(\bar{Y}^*/F(Y^{2*})\bar{F}(\bar{Y}^*)) \neq 0$. But by our assumption $Pr'(\bar{Y}^*/F(Y^{1*})\bar{F}(\bar{Y}^*)) = 0$ and hence (46) would not be satisfied. Thus if the code is lossless then we must have $m = l$. But this forces

$$\prod_{i=1}^{m-1} p_{\bar{x}_i}(\bar{y}_i/x_i^{\,1}) = \prod_{i=1}^{m-1} p_{\bar{x}_i}(\bar{y}_i/x_i^{\,2}),$$

as a small modification of the argument in the preceding paragraph would show. We therefore conclude:

Given any pair F, $\bar{F}$ there is a set $S(F, \bar{F})$ of pairs Y, $\bar{Y}$ such that simultaneously $Pr'(\bar{Y}/F(Y), \bar{F}(\bar{Y})) \neq 0$, $Pr''(Y/F(Y), \bar{F}(\bar{Y})) \neq 0$. Let $y(F, \bar{F}, \bar{Y}^*)$ be the set of Y's such that the pair Y, $\bar{Y}^* \in S(F, \bar{F})$. Let $\bar{y}(F, \bar{F}, Y^*)$ be the set on $\bar{Y}$'s such that the pair Y^*, $\bar{Y} \in S(F, \bar{F})$. And let $\mathfrak{X}(F, \bar{F}, \bar{Y}^*)$ be the set of all $X = F(Y)$, $Y \in y(F, \bar{F}, \bar{Y}^*)$. Then a code will satisfy condition (46) if and only if for all pairs $(F, \bar{F})$ and $\bar{Y}$, $p_{\bar{x}_i}(\bar{y}_i/x_i) =$ constant over all $X \in \mathfrak{X}(F, \bar{F}, \bar{Y})$, where $\bar{X} = \bar{F}(\bar{Y})$.

But if condition (a) of the theorem is satisfied, then above argument together with the necessity that $p_{\bar{x}}(\bar{y}/x) \geqq 0$ and $\sum_{\bar{y}} p_{\bar{x}}(\bar{y}/x) = 1$ shows that $\mathfrak{X}(F, \bar{F}, \bar{Y})$ must consist of either zero (when $\mathcal{S}(F, \bar{F}, \bar{Y})$ is empty) or of one element X. Hence for all F and $Y \in \mathcal{S}(F, \bar{F}, \bar{Y})$ we must have $F(Y) =$ constant.

One way to accomplish above is by letting $F(Y) = \text{const.}$ for all Y, i.e., by associating F with a fixed sequence X. This of course agrees with the assertion of the present theorem. We will show that under the channel conditions (a), (b), and (c) of Theorem 6 there is no other possibility of satisfying (46).

Given the set $S(F, \bar{F})$ of sequence pairs $Y, \bar{Y}$, let us construct non-intersecting subsets $B_i(F, \bar{F})$ as follows: List all sequences Y of $S(F, \bar{F})$ in a column, one sequence to a row. In the next column list in the same way all sequences $\bar{Y}$ of $S(F, \bar{F})$. Next, taking each Y in turn, link it by a straight line to each member of $\mathcal{S}(F, \bar{F}, Y)$ in the second column. Figure 11 illustrates the intended arrangement. Consider the set $C_1(F, \bar{F})$ of all Y and $\bar{Y}$ connected by a succession of links to Y_1, the first sequence in the Y-column, and let $B_1(F, \bar{F})$ be the set of pairs $Y, \bar{Y} \in S(F, \bar{F})$ which can be formed out of elements C_1 (e.g., in Fig. 11 the set $B_1(F, \bar{F}) = \{(Y_1, \bar{Y}_1), (Y_1, \bar{Y}_2), (Y_3, \bar{Y}_1), (Y_3, \bar{Y}_2), (Y_4, \bar{Y}_2)\}$). Next eliminate all the elements of C_1 from the two columns of the diagram and form set C_2 out of all Y and $\bar{Y}$ connected by a succession of links to the first remaining sequence of the Y-column. From C_2 form the set $B_2(F, \bar{F})$ (e.g., in Fig. 11 we have $B_2(F, \bar{F}) = \{(Y_4, \bar{Y}_2)\}$). One can continue in this way forming sets $B_i(F, \bar{F})$ $(i > 2)$ until after the kth step the two-column diagram is empty.

It should be noticed that:

(a) $\bigcup_{i=1}^{k} B_i(F, \bar{F}) = S(F, \bar{F})$.

(b) $B_i \cap B_j = \phi$ if $i \neq j$.

(c) $B_i(F, \bar{F})$ and $B_j(F, \bar{F})$, $j \neq i$, have neither any Y nor any $\bar{Y}$ sequences in common.

We said that for all F and $Y \in \mathcal{S}(F, \bar{F}, \bar{Y})$, $F(Y) = \text{const.}$ But that means that $F(Y) = \text{const.}$ for all $Y \in C_i(F, \bar{F})(i = 1, \cdots, k)$. In

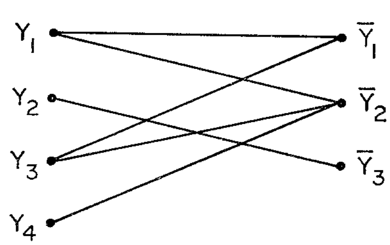

FIG. 11. Sample diagram for determination of sets $B_i(F, \bar{F})$

fact, if $\bar{Y}^1, \bar{Y}^2 \in C_i(F, \bar{F})$, there exists a $Y \in C_i(F, \bar{F})$ such that $(Y, \bar{Y}^1)$ and $(Y, \bar{Y}^2) \in S(F, \bar{F})$. But then $F(Y^*) = \text{const.}$ for all $Y^* \in \mathcal{S}(F, \bar{F}, \bar{Y}^1) \cup \mathcal{S}(F, \bar{F}, \bar{Y}^2)$. Continuing in this way, the assertion follows.

We will next show that because of condition (b) of the theorem, $B_1(F, \bar{F}) = S(F, \bar{F})$, and hence that if (46) is to be satisfied, then $F(Y) = \text{const.}$ for all Y such that $(Y, \bar{Y}) \in S(F, \bar{F})$ for some $\bar{Y}$.

Let $F = f_1, \cdots, f_n$, $\bar{F} = \bar{f}_1, \cdots, \bar{f}_n$ and let the common memory length of the functions f and $\bar{f}$ be m. By convention (5) f_1 and $\bar{f}_1$ map strictly into some signals x_1, $\bar{x}_1$, regardless of any received signals. Let the sets $\mathcal{Y}(x_1, \bar{x}_1)$ and $\bar{\mathcal{Y}}(x_1, \bar{x}_1)$ consist of all signals y_1 and $\bar{y}_1$ such that $\bar{p}_{x_1}(y_1/\bar{x}_1) \neq 0$ and $p_{\bar{x}_i}(\bar{y}_1/x_1) \neq 0$, respectively. Now $f_2(x_1, y_1)$ must be equal to the same signal x_2 for all $y_1 \in \mathcal{Y}(x_1, \bar{x}_1)$. Let $\bar{\mathcal{X}}_2(\bar{f}_2, x_1, \bar{x}_1)$ be the set of $\bar{x}_2$ generated by $f_2(\bar{x}_1, \bar{y}_1)$, $\bar{y}_1 \in \mathcal{Y}(x_1, \bar{x}_2)$. Take any $\bar{x}_2^{\,1}$ and $\bar{x}_2^{\,2} \in \bar{\mathcal{X}}_2(\bar{f}_2, x_1, \bar{x}_1)$. Then, because of condition (b) of the theorem, $\mathcal{Y}(x_2, \bar{x}_2^{\,1}) \cap \mathcal{Y}(x_2, \bar{x}_2^{\,2}) \neq 0$. Hence $f_3(x_1, x_2, y_1, y_2)$ must be equal to the *same* signal x_3 for all $y_1 \in \mathcal{Y}(x_1, \bar{x}_1)$ and $y_2 \in \bigcup_i \mathcal{Y}(x_2, \bar{x}_2^{\,i})$, $\bar{x}_2^{\,i} \in \bar{\mathcal{X}}_2(\bar{f}_2, x_1, \bar{x}_1)$. Continuing this argument for $j = 3, 4, \cdots, n$ we conclude that given any $\bar{F}$, the sequence F must map into the same signal X for all receivable Y.

Finally, we will show that $F(Y)$ must equal a constant for all Y such that $(Y, \bar{Y}) \in \bigcup_{\bar{F} \in \bar{\mathfrak{F}}} S(F, \bar{F})$ for some $\bar{Y}$. Select any pair $Y, \bar{Y}^1$ from the set $S(F, \bar{F}^1)$. If there exists a sequence $\bar{Y}^2$ such that $Y, \bar{Y}^2 \in S(F, \bar{F}^2)$, then $F(Y) = \text{const.}$ for all Y such that $Y, \bar{Y} \in S(F, \bar{F}^1) \cup S(F, \bar{F}^2)$. All we must show then is that for any pair $\bar{F}^1, \bar{F}^2$ and any F, there will exist sequences $Y, \bar{Y}^1, \bar{Y}^2$ such that $Y, \bar{Y}^1 \in S(F, F^1)$, $Y, \bar{Y}^2 \in S(F, \bar{F}^2)$. Select any $Y^1, \bar{Y}^1 \in S(F, \bar{F}^1)$, and any $Y^2, \bar{Y}^2 \in S(F, \bar{F}^2)$. Let $\bar{F}^1(\bar{Y}^1) = \bar{X}^1$, $\bar{F}^2(\bar{Y}^2) = \bar{X}^2$, $F(Y^1) = X^1$. Because of condition (b), there must exist a sequence Y such that $\Pr''(Y/X^1, \bar{X}^1) \neq 0$ and

$$\Pr''(Y/X^1, \bar{X}^2) \neq 0.$$

But then $Y, \bar{Y}^1 \in S(F, \bar{F}^1)$ and $Y, \bar{Y}^2 \in S(F, \bar{F}^1)$ since

$$F(Y) = F(Y^1) = F(Y^2)$$

as we have already shown. This completes the proof of the theorem. Q.E.D.

The conditions (a), (b), and (c) of Theorem 6 are those which a general channel would "naturally" meet. Condition (a) states that there should exist for no $\bar{x}$ any pair of left signals that would be statistically indistinguishable to the right receiver. Condition (b) states that for no left signal x should it be possible for the left receiver to distinguish without any error between two disjoint classes of right signals $\bar{x}$. Finally, condition (c) states that for no signal combination x, $\bar{x}$ should the left received signal be fully determined.

Theorem 6 was proven for arbitrary but fixed n, and for definite choices of code-word sets $\mathfrak{F}$ and $\bar{\mathfrak{F}}$. What statements can be made about coding loss per symbol, $(1/n)E\{I(Y; \bar{Y}/F, \bar{X})\}$ when successive functions f and $\bar{f}$ are selected independently at random with fixed probabilities $P(\cdot)$ and $\bar{P}(\cdot)$?

Let the probability measures $P(\cdot)$ and $\bar{P}(\cdot)$ be given, and denote by $\mathfrak{f}(\bar{\mathfrak{f}})$ the subset of $\{f\}$ $(\{\bar{f}\})$ consisting of those functions $f(\bar{f})$ for which $P(f) \neq 0 (\bar{P}(\bar{f}) \neq 0)$. Let it be possible to construct from elements of $\mathfrak{f}$ a word F^* such that, for all arbitrary but fixed $\bar{F}$ constructed only from elements of $\bar{\mathfrak{f}}$, $F^*(Y_1) \neq F^*(Y_2)$ for some $Y_1, Y_2 \in S(F^*, \bar{F})$. Then if the channel satisfies conditions (a), (b), and (c), the relation $(1/n)E\{I(Y; \bar{Y}/F, \bar{X})\} > \delta > 0$ will hold, where the lower bound δ will be a monotonically increasing function of the probabilities $P(\ )$ of the "offending" symbols f in the sequence F^*. Thus under channel conditions (a), (b), and (c), per symbol losslessness of the left transducer in Fig. 8 is possible only if the collection $\mathfrak{f}$ is equivalent, with respect to the channel and the collection $\bar{\mathfrak{f}}$, to a set of f's such that

$$f(X^m, Y^m) = \phi(X^m)$$

for all Y^m. Assigning low nonzero probabilities to f's such that

$$f(X^m, Y^m) \neq \phi(X^m)$$

will keep the coding loss per symbol down, but on the other hand the average rate $S^n(P(\ ), \bar{P}(\ ))$ (see (30)) would be close to that obtainable by strict encoding of messages into signals. Hence, at least when dealing with channels satisfying conditions (a), (b), and (c) of the preceding theorem, one must accept coding loss as a necessary price for possible improvements over one-way channel type encoding.

V. CONCLUSIONS

A. Optimal Symbol Sources for Lossless Channels

Consider the problem of finding the probability functions $P(f)$ and $\bar{P}(\bar{f})$ which would maximize the weighted sum $S^n + \lambda\bar{S}^n$ (see (30) and (31)) of average information transmitted through the derived two-way channel of Fig. 9 when the signal part of the latter satisfies for both of its directions the condition (a) of Theorem 5 (i.e., when for any signal combination x, $\bar{x}$ the probabilities $\bar{p}_x(y/\bar{x})$ and $p_{\bar{x}}(\bar{y}/x)$ are equal to either zero or one). A simple example is the multiplying channel discussed by Shannon (1961). The alphabets x, $\bar{x}$, y, $\bar{y}$ are all binary and if x and $\bar{x}$ were transmitted, the received signals are determined by the equation $y = \bar{y} = (x, \bar{x})$. Since such channels are lossless then the convenient way is to find the probability functions $q_{X^m Y^m}(x)$ and $\bar{q}_{\bar{X}^m \bar{Y}^m}(\bar{x})$ maximizing $R^n + \lambda \bar{R}^n$ (see (7) and (12)) and then by use of the procedure of Jelinek (1964*b*) find any functions $P(\ )$ and $\bar{P}(\ )$ satisfying equations (27). The latter probability assignment will maximize $S^n + \lambda \bar{S}^n$. As pointed out, the considerable advantage of this approach stems from the fact that maximization over "only" $g(gh)^m + \bar{g}(\bar{g}h)^m$ instead of over $g^{(gh)^m} + \bar{g}^{(\bar{g}h)^m}$ variables is required. Such maximization is still a formidable process, since all attempts to prove a unique local maximum for the function $R^n + \lambda \bar{R}^n$ over the variables q and $\bar{q}$ have so far failed. Moreover, some thought will reveal that the best that can be hoped for in the general case is a theorem proving a unique local maximum for $R^n + \lambda \bar{R}^n$ over the set $\{q\}$ with the set $\{\bar{q}\}$ fixed, and vice versa. It should be noted that the procedure of Jelinek (1964*b*) does not lead to a unique set $P(f)$ for a given set $\{q_{X^m Y^m}(x)\}$, and hence a direct optimization of $S^n + \lambda \bar{S}^n$ would certainly lead to many local maxima.

It was proven in another paper[5] that if and only if the sources in Fig. 8 generate successive symbols independently with probabilities $P(f)$ and $\bar{P}(\bar{f})$, the equivalent signal sources (see Figs. 4 and 6) will generate successive signals with probabilities $q_{X^m Y^m}(x)$ and $\bar{q}_{\bar{X}^m \bar{Y}^m}(\bar{x})$ defined in (27). If the successive symbols generated by the sources in Fig. 8 are in any way dependent, then the equivalent signal sources of Fig. 6 will have infinite memories. It then follows that if for the lossless channel we restrict the sources of Fig. 8 to stationary ones, then we will get close to optimal information transmission through the channel if we make the memory length m of the symbols f and $\bar{f}$ sufficiently large, and if we generate the successive symbols independently. In this way it will be possible to get arbitrarily close to the boundary of the stationary source capacity region—it will never be necessary to employ dependent f- and $\bar{f}$-sources.

It is possible to show (see Jelinek (1962 a), section 7.6) by an argument based on Theorem 6 and on the discussion following its proof that the capacity region G for channels satisfying conditions (a), (b), and (c) of Theorem 6 will be strictly interior to the outer bound G_0 given in Shannon (1961). It should be pointed out in this connection that no two-way channel has yet been constructed for which it could be proven that its capacity region G exceeds the inner bound G_I of Shannon (1961). In fact, it is conjectured in Jelinek (1964 a) that for all symmetrical channels, G is equal to G_I.

B. Optimal Symbol Sources for Lossy Channels

The last statement of Section V.A cannot, unfortunately, be made about optimization of sources in Fig. 8 when the two-way channel allows coding loss. This follows since

$$\begin{aligned} E[I'(Y; F/F)] &+ \lambda E[I''(Y; F/F)] \\ &= \{E[I'(\bar{Y}; X/\bar{X})] + \lambda E[I''(Y; \bar{X}/X)]\} \\ &\quad - \{E[I'(Y; \bar{Y}/F, \bar{X})] + \lambda E[I''(Y; \bar{Y}/X, \bar{F})]\}, \end{aligned} \tag{77}$$

and although independent generation of f and $\bar{f}$ symbols with sufficiently large memory length m will bring the first braced expression on the right hand side to within any desired $\epsilon > 0$ of its maximum, it remains to be shown how the value of the second braced expression is related to dependent and independent generation of successive f and $\bar{f}$ symbols.

Nevertheless, optimization of the left hand expression in (77) over dependent sources is in any "practical" case simply unthinkable, and the mind recoils even at the thought of optimization over independent f and $\bar{f}$ sources. Rather, a quasi-optimization approach suggests itself maximizing the first braced expression on the right hand side of (77) over probabilities $\{q\}$ and $\{\bar{q}\}$, and then finding those probability functions $P(\ )$ and $\bar{P}(\ )$ which would minimize the second braced expression on the right hand side of (77) under the constraint (27). In fact, the latter minimization would again be too complicated, so a further compromise would have to be made, perhaps by modifying the procedure of Jelinek (1964*b*) so as to yield among the possible sets $\{P(f)\}$, $\{\bar{P}(\bar{f})\}$ of size $(g - 1)(gh)^m + 1$ that pair which guarantees a minimum to the second right hand side braced expression in (77).

Unfortunately, it can be shown that the quasi-optimization procedure even without the last mentioned compromise will in general never yield the actual optimum.

Theorem 7. *Given a two-way channel whose both directions satisfy conditions* (a), (b), *and* (c) *of Theorem 6. If the actual capacity region*

G exceeds its inner bound G_I (Shannon, 1961) then for any code-word length n the left hand side of (77) maximized over probability functions $P(\cdot)$ and $\bar{P}(\cdot)$ will always exceed in value the sum of the function

$$E[I'(\bar{Y}; X/\bar{X})] + \lambda E[I''(Y; \bar{X}/X)] \tag{78}$$

maximized over probability functions $\{q_{X^m Y^m}(x)\}$ $\{\bar{q}_{\bar{X}^m \bar{Y}^m}(\bar{x})\}$, with the function

$$-E[I'(Y; \bar{Y}/F, \bar{X})] - \lambda E[I''(Y; \bar{Y}/X, \bar{F})] \tag{79}$$

maximized over probability functions $P(\cdot)$, $\bar{P}(\cdot)$ under the constraint (27) (i.e., *$P(\cdot)$ and $\bar{P}(\cdot)$ are such that the left hand sides of* (27) *consist of the functions $q_{X^m Y^m}(x)$ and $\bar{q}_{\bar{X}^m \bar{Y}^m}(\bar{x})$ which maximized* (78)).

We will omit the proof, as it follows rather closely that of a similar theorem proven for one-way channels with side state information.[9]

Theorem 7 together with the results of Section IV shows that in order to maximize the flow of *useful* information through the channel, one must not in general maximize the *total* flow of information through the channel. Rather, one must make a compromise and send less information through the channel, of which however a greater part is useable for message identification. Thus, up to a certain point, an increase in total information flow through the channel due to strategy coding can be made in such a way that an increase in useful information will correspond to it; beyond that point however, any increase in total information will be accompanied by an even greater increase in coding information loss, so that the net amount of useful information will actually decrease.

Acknowledgment

The author is indebted to his thesis supervisor, Professor R. M. Fano for his encouragement and many helpful suggestions.

Received: December 11, 1962

References

Fano, R. M. (1961), "Transmission of Information." New York.

Jelinek, F. (1962*a*), Coding for discrete memoryless two-way channels. Ph.D. Thesis, Department of Electrical Engineering, M.I.T., January.

Jelinek, F. (1962*b*), Coding for discrete memoryless two-way channels. Quarterly

[9] See Jelinek (1962*a*), Theorem 9-4.

Progress Report No. 66, pp. 239–253. Research Laboratory of Electronics, M.I.T.

JELINEK, F. (1964*a*), Coding for and decomposition of two-way channels. *IRE Trans. Inform. Theory*, January.

JELINEK, F. (1964*b*), An algorithm for construction of probability distributions needed for attainment of capacity in channels with side information. Unpublished observations.

SHANNON, C. E. (1961), Two-way communication channels. *Fourth Berkeley Symp. Probability and Statist.* J. Neyman, ed., Vol. 1, pp. 611–644. Univ. of California Press, Berkeley, California.